We claim:

1. A method of determining the nucleotide sequence at an end of a polynucleotide, the method comprising the steps of:

ligating one or more encoded adaptors to an end of the polynucleotide, each encoded adaptor having an oligonucleotide tag selected from a minimally cross-hybridizing set of oligonucleotides and a protruding strand complementary to a portion of a strand of the polynucleotide; and

identifying one or more nucleotides in each of said portions of the strand of the polynucleotide by specifically hybridizing a tag complement to each oligonucleotide tag of the one or more encoded adaptors ligated thereto.

2. The method of claim 1 wherein said step of ligating includes ligating a plurality of different encoded adaptors to said end of said polynucleotide such that said protruding strands of the plurality of different encoded adapators are complementary to a plurality of different portions of said strand of said polynucleotide such that there is a one-to-one correspondence between said different encoded adaptors and the different portions of said strand.

3. The method of claim 2 wherein said different portions of said strand of said polynucleotide are contiguous.

4. The method of claim 3 wherein said protruding strand of said encoded adaptors contains from 2 to 6 nucleotides and wherein said step of identifying includes specifically hybridizing said tag complements to said oligonucleotide tags such that the identity of each nucleotide in said portions of said polynucleotide is determined successively.

5. The method of claim 4 wherein said step of identifying further includes providing a number of sets of tag complements equivalent to the number of nucleotides to be identified in said portions of said polynucleotide.

6. The method of claim 5 wherein said step of identifying further includes providing said tag complements in each of said sets that are capable of indicating the presence of a predetermined nucleotide by a signal generated by a fluorescent signal generating moiety, there being a different fluorescent signal generating moiety for each kind of nucleotide.

7. The method of claim 5 wherein said oligonucleotide tags of said encoded adaptors are single stranded and said tag complements to said oligonucleotide tags are single stranded such

that specific hybridization between an oligonucleotide tag and its respective tag complement occurs through Watson-Crick base pairing.

8. The method of claim 7 wherein said encoded adaptors have the form:

$$\begin{array}{l} 5'\text{-}p(N)_n(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ \qquad\quad z(N')_r(N')_s(N')_q\text{-}5' \end{array}$$

or

$$\begin{array}{l} \qquad\quad p(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ 3'\text{-}z(N)_n(N')_r(N')_s(N')_q\text{-}5' \end{array}$$

where N is a nucleotide and N' is its complement, p is a phosphate group, z is a 3' hydroxyl or a 3' blocking group, n is an integer between 2 and 6, inclusive, r is an integer between 0 and 18, inclusive, s is an integer which is either between four and six, inclusive, whenever the encoded adaptor has a nuclease recognition site or is 0 whenever there is no nuclease recognition site, q is an integer greater than or equal to 0, and t is an integer greater than or equal to 8.

9. The method of claim 8 wherein r is between 0 and 12, inclusive, t is an interger between 8 and 20, inclusive, and z is a phosphate group.

10. The method of claim 9 wherein members of said minimally cross-hybridizing set differ from every other member by at least six nucleotides.

11. The method of claim 5 wherein said oligonucleotide tags of said encoded adaptors are double stranded and said tag complements to said oligonucleotide tags are single stranded such that specific hybridization between an oligonucleotide tag and its respective tag complement occurs through the formation of a Hoogsteen or reverse Hoogsteen triplex.

12. The method of claim 11 wherein said encoded adaptors have the form:

$$\begin{array}{l} 5'\text{-}p(N)_n(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ \qquad\quad z(N')_r(N')_s(N')_q(N)_t\text{-}5' \end{array}$$

or

$$\begin{array}{l} \qquad\quad p(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ 3'\text{-}z(N)_n(N')_r(N')_s(N')_q(N)_t\text{-}5' \end{array}$$

where N is a nucleotide and N' is its complement, p is a phosphate group, z is a 3' hydroxyl or a 3' blocking group, n is an integer between 2 and 6, inclusive, r is an integer between 0 and

18, inclusive, s is an integer which is either between four and six, inclusive, whenever the encoded adaptor has a nuclease recognition site or is 0 whenever there is no nuclease recognition site, q is an integer greater than or equal to 0, and t is an integer greater than or equal to 8.

13. The method of claim 12 wherein r is between 0 and 12, inclusive, t is an interger between 8 and 24, inclusive, and z is a phosphate group.

14. The method of claim 13 wherein members of said minimally cross-hybridizing set differ from every other member by at least six nucleotides.

15. A method of determining the nucleotide sequences of a plurality of polynucleotides, the method comprising the steps of:

(a) attaching a first oligonucleotide tag from a repertoire of tags to each polynucleotide in a population of polynucleotides such that each first oligonucleotide tag from the repertoire is selected from a first minimally cross-hybridizing set;

(b) sampling the population of polynucleotides to form a sample of polynucleotides such that substantially all different polynucleotides in the sample have different first oligonucleotide tags attached;

(c) sorting the polynucleotides of the sample by specifically hybridizing the first oligonucleotide tags with their respective complements, the respective complements being attached as uniform populations of substantially identical oligonucleotides in spatially discrete regions on the one or more solid phase supports;

(d) ligating one or more encoded adaptors to an end of the polynucleotides in the sample, each encoded adaptor having a second oligonucleotide tag selected from a second minimally cross-hybridizing set and a protruding strand complementary to a protruding strand of a polynucleotide of the population; and

(e) identifying a plurality of nucleotides in said protruding strands of the polynucleotides by specifically hybridizing a tag complement to each second oligonucleotide tag of the one or more encoded adaptors.

16. The method of claim 15 further including the steps of (f) cleaving said encoded adaptors from said polynucleotides with a nuclease having a nuclease recognition site separate from its cleavage site so that a new protruding strand is formed on said end of each of said polynucleotides, and (g) repeating steps (d) through (f).

17. A method of identifying a population of mRNA molecules, the method comprising the steps of:

(a) forming a population of cDNA molecules from the population of mRNA molecules such that each cDNA molecule has a first oligonucleotide tag attached, the first oligonucleotide tags being selected from a first minimally cross-hybridizing set;

(b) sampling the population of cDNA molecules to form a sample of cDNA molecules such that substantially all different cDNA molecules have different first oligonucleotide tags attached;

(c) sorting the cDNA molecules by specifically hybridizing the first oligonucleotide tags with their respective complements, the respective complements being attached as uniform populations of substantially identical complements in spatially discrete regions on one or more solid phase supports;

(d) ligating one or more encoded adaptors to an end of the cDNA molecules in the population, each encoded adaptor having a second oligonucleotide tag selected from a second minimally cross-hybridizing set and a protruding strand complementary to a protruding strand of a cDNA molecule of the sample; and

(e) determining the identity and ordering of a plurality of nucleotides in each of said protruding strands of the cDNA molecules by specifically hybridizing a tag complement to each second oligonucleotide tag of the one or more encoded adaptors;

wherein the population of mRNA molecules is identified by the frequency distribution of the portions of sequences of the cDNA molecules.

18. The method of claim 17 further including the steps of (f) cleaving said encoded adaptors from said polynucleotides with a nuclease having a nuclease recognition site separate from its cleavage site so that a new protruding strand is formed on said end of each of said cDNA molecules, and (g) repeating steps (d) through (f).

19. A method of determining the nucleotide sequence at an end of a polynucleotide, the method comprising the steps of:

(a) ligating an encoded adaptor to an end of the polynucleotide, the encoded adaptor having an oligonucleotide tag selected from a minimally cross-hybridizing set of oligonucleotides and a protruding strand complementary to a portion of a strand of the polynucleotide;

(b) identifying one or more nucleotides in the portion of the strand of the polynucleotide by specifically hybridizing a tag complement to the oligonucleotide tag of the encoded adaptor ligated thereto;

(c) cleaving the encoded adaptor from the end of the polynucleotide with a nuclease having a nuclease recognition site separate from its cleavage site so that a new protruding strand is formed at the end of the polynucleotide; and

(d) repeating steps (a) through (c).

20. The method of claim 19 wherein said protruding strand of said encoded adaptor contains from 2 to 6 nucleotides and wherein step of identifying includes specifically hybridizing successive said tag complements to said oligonucleotide tag such that the identity of each nucleotide in said portion of said polynucleotide is determined successively.

21. The method of claim 20 wherein said step of identifying further includes providing a number of sets of tag complements equivalent to the number of nucleotides to be identified in said portion of said polynucleotide.

22. The method of claim 21 wherein said step of identifying further includes providing said tag complements in each of said sets that are capable of indicating the presence of a predetermined nucleotide by a signal generated by a fluorescent signal generating moiety, there being a different fluorescent signal generating moiety for each kind of nucleotide.

23. The method of claim 22 wherein said oligonucleotide tags of said encoded adaptors are single stranded and said tag complements to said oligonucleotide tags are single stranded such that specific hybridization between an oligonucleotide tag and its respective tag complement occurs through Watson-Crick base pairing.

24. A composition of matter comprising a double stranded oligonucleotide adaptor having the form:

$$\begin{array}{l} 5'\text{-}p(N)_n(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ \qquad\quad z(N')_r(N')_s(N')_q\text{-}5' \end{array}$$

or

$$\begin{array}{l} \qquad\quad p(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ 3'\text{-}z(N)_n(N')_r(N')_s(N')_q\text{-}5' \end{array}$$

where N is a nucleotide and N' is its complement, p is a phosphate group, z is a 3' hydroxyl or a 3' blocking group, n is an integer between 2 and 6, inclusive, r is an integer between 0 and 18, inclusive, s is an integer which is either between four and six, inclusive, whenever the encoded adaptor has a nuclease recognition site or is 0 whenever there is no nuclease recognition site, q is an integer greater than or equal to 0, t is an integer greater than or equal to 8.

25. The composition of claim 24 wherein r is between 0 and 12, inclusive, t is an interger between 8 and 20, inclusive, z is a phosphate group, and said single stranded moiety $(N)_t$ is a member of a minimally cross-hybridizing set.

26. The composition of claim 25 wherein n equals 4 and wherein members of said minimally cross-hybridizing set differ from every other member by at least six nucleotides.

27. A composition of matter comprising a double stranded oligonucleotide adaptor having the form:

$$\begin{array}{l} 5'\text{-}p(N)_n(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ \qquad\quad z(N')_r(N')_s(N')_q(N')_t\text{-}5' \end{array}$$

or

$$\begin{array}{l} \qquad\quad p(N)_r(N)_s(N)_q(N)_t\text{-}3' \\ 3'\text{-}z(N)_n(N')_r(N')_s(N')_q(N')_t\text{-}5' \end{array}$$

where N is a nucleotide and N' is its complement, p is a phosphate group, z is a 3' hydroxyl or a 3' blocking group, n is an integer between 2 and 6, inclusive, r is an integer between 0 and 18, inclusive, s is an integer which is either between four and six, inclusive, whenever the encoded adaptor has a nuclease recognition site or is 0 whenever there is no nuclease recognition site, q is an integer greater than or equal to 0, and t is an integer greater than or equal to 8.

28. The composition of claim 27 wherein r is between 0 and 12, inclusive, t is an interger between 8 and 24, inclusive, z is a phosphate group, and said double stranded moiety

$$\begin{array}{l} -(N)_t \\ -(N')_t \end{array}$$

is a member of a minimally cross-hybridizing set.

29. The composition of claim 28 wherein members of said minimally cross-hybridizing set differ from every other member by at least six nucleotides.